Comparing Spatial and Temporal Variation of Lake-Atmosphere Carbon Dioxide Fluxes Using Multiple Methods

Angela K. Baldocchi¹, David E. Reed¹,², Luke C. Loken¹, Emily H. Stanley³, Hayley Huerd⁴, and Ankur R. Desai¹

¹Department of Atmospheric and Oceanic Sciences, University of Wisconsin-Madison, Madison, WI, USA, ²Environmental Science, University of Science and Arts of Oklahoma, Chickasha, OK, USA, ³Center for Limnology, University of Wisconsin-Madison, Madison, WI, USA, ⁴Department of Environmental Engineering, University of California Merced, Merced, CA, USA

Abstract Lakes emit globally significant amounts of carbon dioxide (CO₂) to the atmosphere, but quantifying these rates for individual lakes is extremely challenging. The exchange of CO₂ across the air-water interface is driven by physical, chemical, and biological processes in both the lake and the atmosphere that vary at multiple spatial and temporal scales. None of the methods we use to estimate CO₂ flux fully capture this heterogeneous gas exchange. Here, we compared concurrent CO₂ flux estimates from a single lake based on commonly used methods. These include floating chambers (FCs), eddy covariance (EC), and two concentration gradient-based methods labeled fixed (F-pCO₂) and spatial (S-pCO₂). At the end of summer, cumulative carbon fluxes were similar between EC, F-pCO₂, and S-pCO₂ methods (−4, −4, and −9.5 gC m⁻²), while methods diverged in directionality of fluxes during the fall turnover period (−50, 43, and 38 gC m⁻²). Collectively, these results highlight the discrepancies among methods and the need to acknowledge the uncertainty when using any of them to approximate this heterogeneous flux.

Plain Language Summary Lakes comprise a small percentage of the landscape, but they are active and complex areas of carbon cycling. Lakes receive mixed carbon inputs from upstream sources, process this carbon internally, store it in sediments and biomass, and export it downstream. In addition, some fraction of the carbon in lakes exchanges into and out of the atmosphere, linking lakes with the global atmosphere. The exchange of carbon dioxide across lake surfaces has globally significant implications, but quantifying these rates has yet to be fully resolved. Here, we compared four methods of estimating diffusive carbon dioxide exchange between the atmosphere and the lake surface. Flux rates generally agreed during the summer, but estimates diverged in the fall, a critical time period with elevated carbon cycling rates. These discrepancies among methods may arise because of the high degree of spatial and temporal variability in gas exchange and our limited ability to portray and scale these processes accurately. In the future, we need to improve both the resolution of observations and how we process those observations to better measure carbon gas exchange between lakes and the atmosphere.

1. Introduction

Lakes are a major component of the Earth’s carbon cycle, and an increasing focus has been placed on carbon dynamics within inland waters (Biddanda, 2017; Tranvik et al., 2009; Williamson et al., 2009). A substantial fraction of the organic carbon that is delivered to or fixed within lakes is outgassed to the atmosphere as carbon dioxide (CO₂) (Cole et al., 2007; Cory et al., 2014). While there is consensus that collectively lakes and other inland waters emit meaningful amounts of CO₂ to the atmosphere, it remains extremely difficult to calculate spatially and temporally resolved emission rates for individual lakes. This difficulty is because the exchange of CO₂ across the air-water interface is driven by multiple physical, chemical, and biological processes in both the lake and the atmosphere that vary at multiple spatial and temporal scales. The scientific community lacks methods to fully capture the spatial and temporal heterogeneity in gas exchange between lakes and the atmosphere. Thus, every estimate of global CO₂ emissions from lakes has uncertainty.
The reason that lake-atmosphere fluxes are difficult to quantify in part is because they vary in magnitude (Raymond et al., 2013), in time (Reed et al., 2018), and across space (Natchimuthu et al., 2016). In many temperate dimictic lakes, seasonal phenologies in ice cover and stratification govern the direction and magnitude of CO2 flux. Large off-gassing events occur during periods of vertical mixing such as ice-off and fall turnover (Denfeld et al., 2016). Lakes with higher productivity show pronounced temporal variation in CO2 flux (Maberly et al., 2012), characterized by influx during the summer periods coinciding with higher rates of primary production (Reed et al., 2018). Thus, for even a single lake, flux estimation needs to be continuous and year-round in order to capture the temporal heterogeneity in gas exchange. Spatially, heterogeneity in metabolic processes, hydrology, and turbulence can have pronounced impacts on CO2 flux from the lake surface. Rivers flowing into lakes typically differ in a number of physical, chemical, and biological properties that can create contrasts in pCO2 in habitats where they enter a lake (Chmiel et al., 2019). Further, spatial heterogeneity varies temporally (Loken, Stanley, et al., 2019; Natchimuthu et al., 2016) due to changes in river flow, lake mixing, and biological processes. Thus, to accurately measure CO2 flux from a single lake, we need to incorporate both spatial and temporal variation.

Any calculations of lake-atmosphere CO2 flux are limited in either spatial or temporal extent. Perhaps the simplest and most cost-effective method for measuring gas efflux from lakes is using floating chambers (FCs) (Bastviken et al., 2015). Chambers are placed atop the lake surface, and the flux is derived from the gas accumulation rate within the chamber. However, flux chambers characterize only a small area of the lake for what is typically a short deployment. Further, the chamber itself can alter turbulence, thus biasing gas exchange within the chamber environment (Vachon et al., 2010). Historically, FCs for CO2 required manual gas sampling followed by laboratory determination of gas concentrations, while newer FCs integrate continuous CO2 sensors and automatic purging mechanisms that allow for longer deployments (Bastviken et al., 2015; Jonsson et al., 2008; Martinsen et al., 2018). While a single measurement is small in its spatial scale, multiple chambers have been used to quantify the spatial variability of gas emissions within and among lake habitats (Natchimuthu et al., 2016; Tangen et al., 2016). Similarly, measuring temporal variability of fluxes using FCs is common but in both cases, characterizing spatial and/or temporal variability with this approach is time intensive. New automated chambers show promise in increasing the duration of continuous observation (Duc et al., 2012).

A common alternative to FCs is modeling exchange rates using the concentration gradient or boundary layer method (F-pCO2) (Cole & Caraco, 1998; MacIntyre et al., 2010; Read et al., 2012). The flux of any gas across the air-water interface is controlled at the molecular level (Kitaigorodskii & Donelan, 1984), and fluxes are estimated using differences between pCO2 on opposing sides of the air-water boundary and an estimate of water turbulence or gas transfer velocity (k). Spatial scales of pCO2 measurements within the water column are on the order of cubic centimeters and typically fixed in space. Estimation of k is typically based on empirically derived models using wind speed, lake size, and/or water density gradients (Crusius & Wanninkhof, 2003; MacIntyre et al., 2010; Read et al., 2012). The difficulty in modeling k stems from the fact that k changes in response to weather events and varies within lakes due to lake and environmental conditions (Natchimuthu et al., 2016; Vachon et al., 2013). Moreover, estimation of k can vary by multiple orders of magnitude simply due to model choice (Dugan et al., 2016). Recent pCO2 studies have shown that scaling k from point measurements to the lake scale strongly underestimates emissions (Mammarella et al., 2015; Schubert et al., 2012). New methods have been developed to quickly quantify spatial variation in pCO2 (Bastviken et al., 2015; Crawford et al., 2015) and have revealed substantial spatial variations in pCO2 and fluxes within individual lakes and reservoirs (Loken, Crawford, et al., 2019; Natchimuthu et al., 2016; Paranaiba et al., 2018). Despite their flaws, boundary layer methods have provided the most frequent and comprehensive understanding of CO2 exchange between lakes and the atmosphere (Balmer & Downing, 2011; Duarte et al., 2008; Raymond et al., 2013), yet most assume spatial homogeneity and are reliant on physical lake models that have large uncertainty.

A third approach for quantifying lake CO2 fluxes is eddy covariance (EC) (Morin et al., 2018; Reed et al., 2018). In contrast to the water-based approaches, EC uses measurements of concentrations of gas in the atmosphere along with high-frequency measurements of wind speeds in three dimensions. While this top-down flux method seems like the silver bullet for quantifying CO2 flux, EC has several assumptions built into estimation and is spatially limited. It relies on measurement during periods with air movement.
sufficient to generate turbulent airflow with eddies and includes uncertainty of footprint models that estimate the area over which fluxes are being measured (i.e., the footprint), with a single flux estimate integrating over ~1 km². Turbulence and footprint issues can lead to upward of 80% of EC data being excluded (Reed et al., 2018). EC estimates represent the average flux from a portion of the lake surface, which bias observations toward nearshore areas (Morin et al., 2018) where most towers are located. Despite these limitations, EC offers a promising method for assessing carbon fluxes from lakes (Vesala et al., 2012).

Because each technique for measuring carbon flux has its limitations, efforts have been made to compare these methods. However, these investigations have been limited to relatively short time periods (Erkkila et al., 2018; Podgrajsek et al., 2016; Schubert et al., 2012). These authors found discrepancies among methods for quantifying CO₂ flux in both space and time. While estimates of carbon fluxes are critical for understanding the global carbon cycle, how best to measure lake-atmosphere fluxes remains challenging and is an open question for the scientific community.

In order to compare methods of quantifying lake-atmosphere fluxes of CO₂, we leveraged multiple concurrent data sets from a single north temperate lake (Lake Mendota, Wisconsin, USA). This lake has been subject to prior CO₂ flux investigations (Loken, Crawford, et al., 2019; Reed et al., 2018). Here, we combined flux records based on measurements of pCO₂ at a moored buoy, measurements distributed across the entire lake surface, EC from a tower located at the end of a narrow peninsula, and FC. The overarching question of this work is: Are lake-atmosphere CO₂ flux estimates consistent among pCO₂, FC, and EC methods? Due to multiple temporal and spatial scales which the independent observations are taken over, we seek to answer the question using (1) analysis of flux distribution over multiple seasons, (2) quantifying cumulative sums of carbon flux, (3) direct comparison of methods, and (4) spectral time series analysis of fluxes.

2. Methods

2.1. Site Description

Lake Mendota is a well-studied lake located in Southern Wisconsin, USA (43.1°N, 89.4°W) and is part of the North Temperate Lakes Long-Term Ecological Research (NTL-LTER) program. It is dimictic and eutrophic, with a surface area of 39.9 km² and a maximum depth of 25.3 m (mean 12.7 m). The majority of the lake’s watershed is composed of agricultural and urban land uses, resulting in elevated nutrient concentrations and high productivity (Carpenter et al., 2007). Thermal stratification typically occurs between May and October and ice cover from late December through March. We defined seasons using water column temperature gradients with spring and fall as periods in which the water column was isothermal, while in summer the lake was thermally stratified.

2.2. Flux Estimates

2.2.1. Fixed Point Concentration Gradient Method (F-pCO₂)

Since 2006, NTL-LTER has managed a monitoring buoy on Lake Mendota that is moored above the lake’s deepest point (43.0995°N, 89.4045°W). The buoy is equipped with meteorological and limnological sensors and is deployed seasonally (approximately April through October), capturing the majority of the ice-free season. In 2015, a Turner Designs C-sense pCO₂ sensor (Turner Designs, San Jose, USA; 0- to 4,000-ppm range, 3% accuracy, ±120 ppm) was added to the buoy and installed at 0.5-m depth. For this study, we used wind speed, surface water temperature, and surface pCO₂ (Magnuson et al., 2019). Wind speed was measured at a height of 2.7 m above the lake surface using an anemometer (R. M. Young Marine Wind Monitor). Water temperature and pCO₂ were measured at a depth of 0.5 m using a RBR concerto thermistor string and a Turner C-Sense CO₂ sonde, respectively. Wind speed and water temperature were measured every 30 min, while pCO₂ was measured every 15 min. pCO₂ in air was measured from an in situ spectroscopy gas analyzer (Picarro, inc. G2401 Gas Concentration Analyzer) located at a nearby building.

Using data collected at the buoy, we calculated the diffusive efflux of CO₂ from the lake surface to the atmosphere according to:

\[
 Flux = k_{gas} \times kh \times (pCO_{2water} - pCO_{2air})
\]  

This fixed-point boundary layer method (F-pCO₂) is based on the partial pressure gradient between the water (pCO₂water) and the atmosphere (pCO₂air). Multiplying this difference by Henry’s law constant
(kh) converts to molar units and by the gas transfer velocity (k_gas) to generate diffusive flux estimates. We estimated k_gas using concurrent wind speed and water temperature recorded at the buoy following Weyhenmeyer et al. (2012), applying the k_goo lake area model and Schmidt model coefficients provided as Model B in Raymond et al. (2013). Henry’s law constant (kh) was calculated using atmospheric pressure and temperature dependence models provided in Plummer and Busenberg (1982). pCO_2 flux estimates were computed at 30-min intervals. To temporally match observations between methods, a subset of F-pCO_2 was used from 8 a.m. to 12:00 p.m., the time period that overlapped with the majority (>90%) of the spatially explicit pCO_2 sampling times (described below).

2.2.2. Spatial Concentration Gradient Method (S-pCO_2)
In addition to the F-pCO_2-based flux estimation at the buoy, we also compared flux estimates using pCO_2 measurements from the entire lake surface (S-pCO_2). For the entire ice-free period of 2016, Loken, Crawford, et al. (2019) generated CO_2 efflux estimates at 988 points distributed in a gridded pattern across the lake surface. Efflux estimates were based on measurements of pCO_2 collected using a boat-mounted water sampling system. Loken, Stanley, et al. (2019) configured a motorboat with water pumps, tubing, a gas equilibrator, a GPS, and water sensors (including a Los Gatos Research Ultraportable Greenhouse Gas Analyzer) to continuously measure (1 Hz) the water surface as the boat traveled across the lake. Bubble free water is extracted from below and in front of any prop created turbulence, with little disturbance to either concentration or flux measurement. S-pCO_2 measurements had an accuracy of ±0.3 ppm. On 26 sampling days spanning the entire ice-free period, ~10,000 pCO_2 measurements were collected over an ~3-hr window in the morning. They used the point measurements to interpolate pCO_2 across the lake surface and, similar to the F-pCO_2 method, calculated efflux using the difference in pCO_2 between the water and the air. To match the spatial pCO_2 data set, Loken, Crawford, et al. (2019) used a spatially explicit k model (Vachon et al., 2013), which takes into account wind speed and direction and allows k to vary across the lake surface. Daily pCO_2 at each of the 988 points were estimated by temporal interpolation, which they combined with daily spatially explicit k estimates to calculate daily efflux. Two subsets of S-pCO_2 data were used to quantify spatial variability, 10 stratified random points from the entire lake and S-pCO_2 measurement locations from within the EC footprint.

Both F-pCO_2 and S-pCO_2 methods depend on wind speed for accurate k estimates. Wind data for this work have an accuracy of 0.3 m s^{-1}. k accuracy varies in time and space but is approximated to be 0.6 cm hr^{-1}. With both methods using the same wind speed for their k estimates, there is a small degree of dependence between the methods.

2.2.3. Flux Chamber Diffusion Method (FC)
We conducted four FC campaigns between 6 July 2017 and 24 April 2018. CO_2 sensors (Sensair K30, ±30 ppm) were installed inside floating plastic chambers with a foam collar of diameter 0.3 m and a height of 0.12 m. Flux rates were calculated using the chamber dimensions (surface area and volume) and continuous pCO_2 measurements within the enclosed headspace. Each 24-hr sampling campaign consisted of seven sampling trips spaced every 4 hr with the goal of measuring flux rates over a complete diel cycle. For each measurement, we placed two chambers on the lake surface in the middle of the lake (same location as the buoy) and let them drift for 5 min. We repeated the FC procedure three times per chamber and calculated the average of the six flux measurements. CO_2 flux was calculated as follows:

\[ \text{Flux} = \frac{\Delta pCO_2}{\Delta t} \times \frac{V}{SA} \]  

(2)

where V is the chamber volume (0.03114 m^3), SA is the chamber bottom area (0.071 m^2), and t is time (s). Prior to the first campaign, we calibrated all sensors using N_2 gas and the “zero calibration” method per Bastviken et al. (2015). For all subsequent campaigns we re-confirmed the zero CO_2 readings using N_2 gas.

2.2.4. EC Method
EC flux observations (Ameriflux site: US-PnP, doi: 10.17190/AMF/1433376) were collected from a tower at the end of an ~50-m-wide peninsula on the shore of Lake Mendota (Figure 1) starting on 20 June 2016. These flux observations were made with a sonic anemometer (CSAT3, Campbell Scientific, Logan UT, USA) and open-path infrared gas analyzer for CO_2 and water vapor gas concentration (LI-7500A, Li-Cor, Lincoln,
NE, USA, ±5 ppm) at a height of 12.4 m above the lake on a 0.95‐m boom, along with measurements of air temperature and humidity (Vaisala, Inc. HMP45C). Measurements of incoming solar radiation and atmospheric pressure were collected from a nearby meteorological tower located on the roof of the Atmospheric, Oceanic, and Space Sciences building at the University of Wisconsin.

Eddy fluxes were calculated based on the covariance of vertical wind velocity and scalar concentrations following the approach of Mauder and Foken (2015), with quality control flags for stationarity, integral turbulence, and propagate estimates of random error. Typical corrections were applied, including planar fit rotation, and Webb-Pearson-Leuning density corrections, except for \( u^* \) filtering. A change point detection method is not possible, given the dependence of flux on \( u^* \) and instead we use a fixed value for cutoff (0.1 m s\(^{-1}\)). Gap filling was performed using the Marginal Disturbance Smapling (MDS) method with REddyProc (Reichstein et al., 2005), with uncertainty estimated using both random uncertainty methods for turbulent flux sampling (Salesky et al., 2012) and gap filling (Moffat et al., 2007). Using an eddy flux surface flux footprint model (Kljun et al., 2015), we identified and removed non-lake data at 30‐min timescales, primarily when winds were from the forested portion of the peninsula. After footprint screening and quality control, 26% of data were retained. Lacking a concentration gradient profile, below-sensor storage fluxes were not measured and are assumed to average to zero at the daily scale (Xu et al., 2019), which is common at most sites.

### 2.3. Comparison of Methods

Flux estimates varied in temporal and spatial coverage (Table 1). EC-based fluxes were collected continuously since 2016. Buoy-based F-pCO2 estimates are also continuous since this time, with the exception of winter months. We only have S-pCO2 rates for the ice-free period of 2016, which were collected approximately weekly and daily rates that were modeled by interpolating \( \rho \text{CO}_2 \) through time.
Table 1

<table>
<thead>
<tr>
<th>Method</th>
<th>Measurement period</th>
<th>Sampling frequency</th>
<th>Spatial extent</th>
<th>Spatial resolution</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed point concentration</td>
<td>Open water seasons (approx. April–October) 2016–2018</td>
<td>15 min</td>
<td>Single point</td>
<td>10 cm²</td>
<td>Magnuson et al. (2019)</td>
</tr>
<tr>
<td>gradient (F-pCO₂)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial concentration gradient (S-pCO₂)</td>
<td>March–December 2016</td>
<td>14 days</td>
<td>Whole lake</td>
<td>200 m²</td>
<td>Loken, Stanley, et al. (2019)</td>
</tr>
<tr>
<td>Flux chamber diffusion (FC)</td>
<td>Four measurement campaigns, July 2017 to April 2018</td>
<td>5-min sampling, every 4 hr for 24 hr</td>
<td>Single point</td>
<td>0.28 m²</td>
<td>A. Desai (2019)</td>
</tr>
<tr>
<td>Eddy covariance (EC)</td>
<td>June 2016 to August 2018</td>
<td>30 min</td>
<td>1 km²</td>
<td>1 km²</td>
<td>A. Desai (2018)</td>
</tr>
</tbody>
</table>

Because of varying temporal resolution among data sets, we converted all data sets to daily averages, representing the coarsest temporal scale. Using the S-pCO₂ flux estimates, we generated two additional spatial data sets. First, we randomly selected 10 stratified points from the entire lake to visualize spatial variability across the lake (S-pCO₂ Stratified Points). Second, we subset the S-pCO₂ data set by only including flux estimates from within the EC footprint (S-pCO₂ Tower Footprint) for a comparison between these two methods that was not confounded by differences in sampling areas. Cumulative fluxes from 2016 were calculated from F-pCO₂, S-pCO₂, and EC observations. For the entire study period, the S-pCO₂ data had 17% of data functionally usable (26 sampling days over the growing season), F-pCO₂ had 86% data retained, and as previously mentioned, 26% of EC data were retained.

In addition to comparing similarity in seasonal pattern and magnitude, we also wanted to determine if the different methods exhibited similar temporal variance. To do so, we calculated a fast Fourier power spectrum for EC, S-pCO₂, and F-pCO₂ gap-filled daily net ecosystem exchange. Buoy winter fluxes were assumed to be zero for the purpose of this analysis. Data analysis was done in Matlab R2019a and IDL 8.6.0.

3. Results

3.1. Patterns Among Methods

Footprint modeling revealed that the EC footprint originated primarily from open water, with very little apparent input from the terrestrial peninsula (Figure 2a), with the distance of maximum contribution of fluxes on average being 40 m, while the distance containing 80% of flux contribution was 410 m. Friction velocity (u*) values were high due to winds crossing the peninsula, showing increased turbulence due to the tree canopy (Figure 2b). While winds originated from all directions, wind speeds were lower over the peninsula as sampling sites varied among methods. Both the FC for the two methods (see Loken, Crawford, et al., 2019 for details). Thus, these three data sources (S-pCO₂, F-pCO₂, and EC) overlapped from June to December 2016. We collected FC flux rates seasonally starting in summer 2017 (28–29 July 2017, 28–29 October 2017, and 23–24 April 2018). In addition to temporal overlap, we must also consider spatial coverage as sampling sites varied among methods. Both the FC- and F-pCO₂-based rates were determined at the center of the lake. The EC rates reflect the area surrounding the tower along the lake's southern shoreline, and S-pCO₂ covered the entire lake surface (Figure 1).

In all years, F-pCO₂ flux estimates followed a similar pattern of near zero or slightly negative fluxes denoting CO₂ movement from the atmosphere to the lake during spring and summer months before becoming strongly positive (net CO₂ efflux from the lake to the atmosphere) in the fall (Figures 3 and 4). Daily-averaged fluxes varied from −1.2 to 4.1 μm m⁻² s⁻¹ across all dates with a SD of 0.62. This same pattern was also demonstrated by the S-pCO₂ method (Figures 3 and 4), and flux estimates were similar in magnitude and direction as the F-pCO₂ results in 2016 (−0.39 to 1.6 μm m⁻² s⁻¹, SD of 0.36). The limited set of FC deployments also followed the same general pattern of CO₂ influx to the lake in spring, a weaker influx during summer, and efflux in the fall (Figures 3b–3d). However, the range of FC flux values was wider than for the two pCO₂-based methods (−22.5 to 18.1 μm m⁻² s⁻¹, CV of 2.51).
Fluxes derived from the EC method were characterized by higher variation, often shifting from negative to positive fluxes within a period of 1–3 days. Daily-averaged fluxes varied from $-22.5$ to $18 \mu \text{mol m}^{-2} \text{s}^{-1}$, and the coefficient of variation was 3.13. There were no clear seasonal patterns in terms of magnitude, direction, or variance although large CO$_2$ uptakes were recorded prior to ice-on in both 2016 and 2017, and negative and smaller positive fluxes were more common during ice-covered winter days.

3.2. Comparisons Among Methods

Differences among methods were clearly illustrated when flux data were expressed as cumulative flux (Figure 5). All methods indicated that the lake was a slight CO$_2$ sink over the summer; however, estimates diverged substantially during fall. Both the S-pCO$_2$ and F-pCO$_2$ methods consistently indicated CO$_2$ flux into...
the lake all summer and substantial CO₂ flux out of the lake during fall. At the end of the year, the cumulative flux based on F-pCO₂ was 15% higher (43.4 vs. 37.7 gC m⁻²) than flux based on S-pCO₂, but both followed similar temporal trends. In contrast, the EC method suggested that the lake fluctuated between CO₂ source and sink behavior with a high degree of variability on the weekly timescale. At the end of summer (day ~268), the EC-based cumulative flux was comparable to the boundary layer-based rates. However, during fall, once mixing begins, the EC cumulative flux became progressively more negative, suggesting that the lake became a more substantial CO₂ sink.

CO₂ fluxes based on FC (flux chamber) agreed in magnitude and direction with the F-pCO₂ during spring, summer, and fall (Figures 3b–3d). Comparing FC with EC, the two methods disagreed in flux magnitude during summer and direction during fall. The discrepancy between methods could be caused by the temporal or spatial resolution of observations. The daytime EC data more closely aligned with the F-pCO₂ and S-pCO₂ result during the summer. These methods agreed that the daytime flux of CO₂ during the summer was consistently into the lake. During the fall, the daytime EC fluxes remained negative, suggesting a consistent flux of CO₂ into the lake. Spatially, the S-pCO₂ results within the EC footprint were consistent with the majority of the S-pCO₂ data. This suggests the lake was relatively homogeneous in regard to flux rates, with subset S-pCO₂ locations showing ~20% variability in accumulated fluxes at the end of the year. Temporal subsets of EC data show differences during the summer with the full day EC data but ultimately small differences in accumulated fluxes at the end of the year. Average EC error was 38.9%, with larger accumulated errors during the fall.

Directly comparing estimates using linear regression models further demonstrates the dissimilarity among methods. The two concentration gradient methods, F-pCO₂ and S-pCO₂, agreed in magnitude and direction ($R^2 = 0.58$, $p$ value < 0.001; Figure 6a). When flux estimates were categorized by season, data from the summer were tightly clustered, while data from the fall were more scattered. Comparing EC to S-pCO₂ (Figure 6b), there was poor agreement ($R^2 = 0.07$; $p = 0.03$), and the regression model had a negative slope. Thus, daily flux rates using EC disagreed in direction with the concentration-based methods.
Fourier power spectral decomposition (Figure 7) of daily flux from EC, F-\(p\)CO\(_2\), and S-\(p\)CO\(_2\) data all had similar patterns over the quantifiable frequencies, with highest spectral power seen in EC time series, S-\(p\)CO\(_2\), and finally F-\(p\)CO\(_2\). Seasonal and synoptic (3–10 days) variability dominate all three, though the EC tower also shows a sub-monthly (~20 day) mode of variability not seen in the other two.

4. Discussion

Few studies have used multiple measurements of multi-year lake-atmosphere fluxes to address systemic biases in methods. Using concurrent multi-year records from a single lake, we showed divergent behavior among flux estimates, particularly during the fall turnover period. EC-based calculations had large and opposing sign CO\(_2\) flux estimates compared to FC and concentration gradient-based methods (F-\(p\)CO\(_2\) and S-\(p\)CO\(_2\)). FC-based methods agreed in direction and magnitude as \(p\)CO\(_2\)-based methods; however, we lack sufficient FC coverage to interrogate the validity of this agreement. Together, these results suggest that at least for this lake and these estimates, EC and concentration gradient methods for estimating CO\(_2\) flux differ dramatically.

The spatial and buoy-based concentration gradient estimates closely agreed. Both estimates followed similar seasonal patterns, indicating
that the lake was taking in CO$_2$ from the atmosphere during the summer and emitted a substantial amount during the fall. The buoy-based data showed this seasonal phenology in three consecutive years (Reed et al., 2018), aligning with other studies of productive lakes (Maberly, 1996) and the perception that productive lakes behave as CO$_2$ sinks during the summer (Balmer & Downing, 2011). The agreement between the spatial and buoy-based concentration data suggests low spatial heterogeneity in CO$_2$ fluxes across the surface of Lake Mendota. On average most of the lake surface was within a 0.2 $\mu$mol m$^{-2}$ s$^{-1}$ range in CO$_2$ flux (Loken, Crawford, et al., 2019). Low spatial heterogeneity in CO$_2$ concentration may reflect the lake’s high buffering capacity, which would dampen variation in CO$_2$ caused by spatial heterogeneity in metabolic processes (Loken, Crawford, et al., 2019). Spatial variability in CO$_2$ flux is small compared to the seasonal variability from all our CO$_2$ flux methods (Figures 3 and 5). However, spatial heterogeneity increased during fall turnover, making the buoy location less representative of the whole lake during this period (Loken, Crawford, et al., 2019). During periods of chaotic water mixing, the representativeness of a single location decreases (Erkkila et al., 2018). Thus, we suspect that the discrepancies among methodologies during the summer season are not due to spatial heterogeneity in gas exchange across the lake surface.

With a limited number of FC observations, FC data approximately matched F-pCO$_2$ and S-pCO$_2$ during the spring and summer. Comparing FC and F-pCO$_2$ methods, López Bellido et al. (2009) and Vachon et al. (2010) found that FCs were systematically higher than F-pCO$_2$, due to site- and time-specific gas transfer velocities and chamber effects on water turbulence. They used daily concentration measurements and hence were not able to access daily patterns. Podgrajsek et al. (2014)

\begin{figure}[h]
\centering
\includegraphics[width=\linewidth]{figure6}
\caption{Daily mean S-pCO fluxes versus F-pCO$_2$ (a) and EC (b) and F-pCO$_2$ versus EC (c). Summer data are plotted as open circles, fall data as *. Linear regression line (dashed) and one-to-one line (dotted). Statistics ($p$ and $R^2$) for linear regression included.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\linewidth]{figure7}
\caption{Fourier power spectral decomposition of daily EC (green), F-pCO$_2$ (purple), and S-pCO$_2$ (orange) CO$_2$ flux.}
\end{figure}
found that FC and EC fluxes generally agreed, except when $pCO_2$ varied within the EC footprint. This is expanded on to show higher EC-derived $CO_2$ fluxes at night relative to F-$pCO_2$ and also that F-$pCO_2$ methods need to account for convection within the water column (Podgrajsek et al., 2016). Both of these studies show that the high degree of spatial and temporal variability of water-atmosphere fluxes and differences between methods, in their case, could be explained by differences in measurement footprint areas. Erkkila et al. (2018) found that F-$pCO_2$-based estimates were lower than EC, while those based on FC were higher than EC estimates. Together, there does not appear to be an emerging trend among results, other than EC fluxes can be typically higher at night, hinting at possibility of convective gas transfer or some sort of decoupling of lake surface to EC measurement height.

$k$ may be responsible for the discrepancy among flux estimates. The $k$ model underlying our concentration gradient-based methods may have not adequately portrayed turbulence at the lake surface. Convective mixing within the water column introduces error into F-$pCO_2$ methods (Podgrajsek et al., 2016). Our models base $k$ on wind speed, but the effects of individual wind events on lakes are highly variable. For example, 2 days with similar wind speed and direction likely does not have the identical patterns of surface turbulence and our models lack resolution at fine spatial scales (Loken, Crawford, et al., 2019). While only using short periods (1–3 days), Eugster et al. (2003) used EC and chambers from Alaska and Switzerland to show the importance of convective mixing due to lake-atmosphere fluxes, with significant differences between methods during periods of stratification and with deep, penetrative convection. Ultimately, concentration gradient-based estimates rely on measurements of wind and/or water density collected at a single location to represent $k$ over the whole lake, adding uncertainty to gas efflux estimates. Here, both concentration gradient fluxes depend on the same wind speed measurement, removing true independence between the two.

Variations in lake mixing depth and potential periods of stratification could explain differences. Here, with a relatively large and windy lake, this seems unlikely. During periods of higher stratification, methods typically agree. Only during the fall, when there is full-lake mixing do we see the large differences between the EC and gradient methods. So while this could be one source of error, it would appear to be limited in absolute size.

Another possible source of uncertainty is rapid buffering in the upper layer of the water column. Any potential buffering would be unlikely to persist over time, and that bias should be predictable based on mixed (e.g., wind speed) or alkalinity lake conditions. The S-$pCO_2$ data are sampled close to the surface, where buffering would be limited, and with the good agreement between S-$pCO_2$ and F-$pCO_2$ methods, the possibility is unlikely.

Buoyancy-driven turbulence is more important than wind-driven turbulence in smaller lakes (Read et al., 2012) and at night (Podgrajsek et al., 2014). While both wind-driven and convective turbulence play roles in gas exchange in Lake Mendota, wind shear is likely a bigger factor in this present study due to the size of the lake. With a large fetch and relatively steady winds, we do not expect convection to be a major driver of gas exchange in Lake Mendota. Comparing $k$ among lakes, smaller lakes have a bigger contribution from convective mixing, while for larger lakes Read et al. (2012) argue that wind is more important. However, the influence of convective mixing may be more important near shore, at night, and during fall turnover. All three factors may provide some clues why the EC estimate diverges so dramatically from the other methods. The EC tower is located near two relatively shallow bays, where thermal convection may cause an increase in $k$ and flux at night. Further, these bays may have irregular currents and breaking waves, which may elicit further enhancement of flux rates within the tower footprint. How these periods align with $CO_2$ saturation would determine the actual flux rates in the footprint. Modeling water currents and convective mixing within the tower footprint is challenging and would require a high-resolution whole lake three-dimensional model. We currently lack data at fine-scale resolution in lakes to fully model $k$ in lakes, which remains an open challenge. While the EC data seem like a promising tool to capture this fine-scale variation in gas exchange, additional methodology developments are needed.

With an increase in the availability of oxygen data and the derived oxygen flux estimates, potential errors in carbon flux measurements can be highlighted by dissolved oxygen (DO). Reed et al. (2018) showed DO fluxes to be more variable than EC $CO_2$ fluxes but largely agreeing over the summer and fall seasons. Using EC measured $CO_2$ fluxes and DO data, Morin et al. (2018) found microbial activity and DO to be more connected to $CO_2$ outgassing, relative to net $CO_2$ exchange. Timescales for these works are often large, and more
research on connecting DO and EC fluxes in smaller timescales could show additional light on net CO₂ exchange.

Another possible explanation is potential biases in EC measurements during periods of low turbulence, complex turbulence, or advection in the atmosphere. Morin et al. (2018) noted in a model study the role of tower height and lake-land circulations in driving eddy transport that would bias traditional flux calculation based on half-hourly Reynold's decomposition. As the surface cools, enhanced low-level atmospheric stability may suppress turbulence, leading to larger than typical storage or advective contribution to surface fluxes (Lee et al., 2004). As noted by Xu et al. (2019), below-sensor storage flux calculation can be critical to correcting tower-measured flux to represent surface flux, especially periods around sunrise and sunset. However, while we lack storage flux observations at this site or models of local circulation and turbulence on the peninsula, there is no evidence in the data of a preferential circulation during fall or other periods of stable conditions. Further work on data quality filtering of EC is necessary to build confidence in its use over lakes.

EC may have other benefits, even when subject to potential systematic bias. Here, when examining the spectral density of the multiple observations, the EC observations show a 20- to 30-day frequency not observed by the other methods, including the similarly high-frequency buoy measurements. Eugster et al. (2003) also conclude that EC methods should be used in order to collect process-scale data from the full season. Similarly, Podgrajsek et al. (2016) suggest that the high temporal resolution of EC is crucial to resolve diel changes in flux, combined with measurements within the water column with high (30 min) frequency. Reed et al. (2018) used a different EC observation data set on Lake Mendota, not used here due to a large amount of gaps from that tower's location during the study period, which showed high degrees of coherence between CO₂ flux and air temperature at a similar sub-monthly (20–30 days) timescale. An emerging trend in aquatic flux literature explores this monthly timescale of variation where Liu et al. (2011) and Liu et al. (2016) connect synoptic weather patterns to mixing, and Shao et al. (2015) and Ouyang et al. (2017) show monthly correlation between CO₂ flux and chlorophyll and algal blooms.

There are ways to capture this 20- to 30-day timescale without high temporal coverage. Previously, Natchimuthu et al. (2016) used a multi-year FC data set and then sub-sampling the observations following the methods of Wik et al. (2016). They concluded that only ≥8 measurement days, distributed over multiple seasons, and high enough spatial coverage (≥8 locations during summer, ≥5 during spring and fall) are key for representative (± 20%) flux estimates at the annual timescale. However, they note that the flux estimates would be biased if observations excluded episodic events such as lake circulation patterns, diel or seasonal variation, or high flux areas from a lake. There is a mismatch between what the EC literature is concluding about needing high temporal resolution observations and the FC literature about only needing ≥8 days for CO₂ flux estimates (Natchimuthu et al., 2016). We argue that while it may be possible to estimate annual fluxes from a small number of sample days, functionally, we think it would be difficult to observe only 8 days of FC fluxes and have a high degree of confidence that we have captured the temporal processes needed. Ultimately, we do judge the flux signal found at the 20- to 30-day frequency as important and the best way to capture appears to be EC methods.

5. Conclusions

While major advances have been made, quantifying lake-atmosphere fluxes from individual lakes over multiple spatial and temporal scales remains a challenge. Lakes are an important factor in carbon cycling at both global and local scales. Accurately accounting for temporally and spatially heterogeneity in the flux of carbon across lake surfaces is vital for incorporation and constraining process-based predictions within lake models.

Overall, there is a need for increased spatiotemporal resolution in studies of CO₂ exchange between lakes and the atmosphere. Multi-year temporal data collection is essential to capture, diel, monthly, and seasonal patterns. Spatially, there is still an open question as to which method is capturing flux magnitude correctly, as each method integrates different processes into the observation. This is done most explicitly when choosing between multiple k models but is also implicated when screening EC data. There is no emerging trend in magnitude or direction between methods, and additional work is needed to bridge spatiotemporal scales.
Conflict of Interests

The authors declare no competing financial interests.

Data Availability Statement

Data for eddy covariance (US-PnP) can be found at Ameriflux (https://doi.org/10.17190/AMF/1433376). CO2 concentrations used in the spatial (data set 337, https://doi.org/10.6073/pasta/fe9c5437f67254f521bf5f7e0308bf93) and temporal concentration gradient (data set ID 129, https://doi.org/10.6073/pasta/9bc6d26f81aa30f05f73766ca0410b) can be found at the NTL-LTER database (https://liter.limnology.wisc.edu/data) and are indexed in the Environmental Data Initiative. Floating chamber data have been deposited into the Environmental Data Initiative database (https://doi.org/10.6073/pasta/fe6a915989753aba6f1886b095e7a52d0).

Acknowledgments

This study was supported by the National Science Foundation (NSF) Atmospheric and Geospace Sciences Postdoctoral Fellowship Program (GEO-1403096) and the NSF Long-Term Ecological Research (LTER) program award to North Temperate Lakes (NTL) (DEB-1440297). We would like to thank Jonathan Thom for field work help, Yost R. for keeping momentum going on the project, and John Maginnis for his perseverance and assistance on the FC 2017-18 campaigns. We thank Mark Johnson and one anonymous reviewer for their thoughtful peer review comments.

References


10.1029/2019JG005623


